Submitted to the 2001 Symposium on Document Image Understanding Technology
April 23-25, 2001, Columbia, MD

Digital Camera for Document Acquisition

Francis (Pete) Fisher
U.S. Army Research Laboratory
(pfisher@arl.army.mil)

Abstract

A variety of military operations utilize information collected from documents in the field. These documents, collected or captured through various means, may be in foreign languages and encompass a wide range of document types and sizes. Army Research Laboratory (ARL) has developed the FALCon system (Forward Area Language Converter) to permit non-linguists to assist translators and linguists by triaging foreign language documents and prioritizing those documents for translation and evaluation. One difficulty that users reported in pilot field tests of FALCon was that the sheet-fed scanner incorporated in the FALCon system was not suitable for certain document types. Documents that were very small, stapled or bound, or printed on stiff or poor quality paper could not be scanned into the FALCon system. In order to expand the types of documents that can be processed using FALCon, ARL is evaluating commercial digital cameras as a possible replacement for the sheet-fed scanner. Document images captured using the digital camera are passed through the FALCon process in a manner similar to that for scanned document images. We are evaluating digital cameras with respect to document imaging capability, ease of use, ease of image transfer, and perceived survivability in field environments. This paper will describe our digital camera evaluation, the digital camera selected for integration with FALCon, the integration of the digital camera into the FALCon system, and the final system capabilities.

1 Introduction

During military actions on foreign soil soldiers can capture large quantities of foreign language documents. Most soldiers involved in such operations are not likely to be able to read and understand these documents. The military maintains linguists, each trained in one or more languages, to evaluate these captured documents. The problem that can arise, particularly for operations in urban environments, is that solders in the field can capture documents in much greater quantities than can be evaluated by the limited number of linguists that are available. Army Research Laboratory (ARL), in conjunction with other military and government agencies, has developed a portable system called FALCon (Forward Area Language Converter) to assist soldiers in the field with the evaluation and triage of foreign language documents.

The FALCon system, which can be operated by non-linguists, provides the user with an English conversion of foreign language documents and an automated key word search. Users can quickly evaluate the intent of a document that they originally could not read, permitting them to support linguists in the field by prioritizing those documents for translation and evaluation. This permits the limited resources of the linguists to be focused on documents deemed most important to the mission.

Figure 1 shows a block diagram of the FALCon process. The FALCon system consists primarily of a scanner and personal computer (PC) with four software modules. The scanner software stores document images and permits users to edit those document images to remove unwanted content. Users select the image that they want converted and then click the FALCon button to start the process. The scanner software then passes the document image to the OCR (optical character recognition) software where it is converted from an image file to a foreign language text file. The OCR software then passes the foreign language text file to the MT (machine translation) software where the foreign language text is converted to text in English. The resulting English text is then scanned for keywords in order to measure the relevance of the document to the specified keyword list. At the conclusion of the process the PC displays a window showing the foreign language text (OCR output), a window showing the English text (MT output) with found keywords marked in red, and a window showing the number of keywords found. Users can then browse the English result and check the keywords in order to access the importance of the document. If desired the user can then save files from any or all of the process steps for further evaluation by a trained linguist. A user interface software module is included to simplify the process of setting language parameters in the three primary software modules. Because users may already have a PC at their location, FALCon is also available as software only. This permits FALCon to be added to a wide variety of systems for evaluation and use by military personnel.

As part of the initial integration effort, ARL provided seven prototype systems to Army users for pilot field-testing in Bosnia in 1997. Feedback from this pilot field test included several reports of documents that could not be evaluated using the FALCon system because

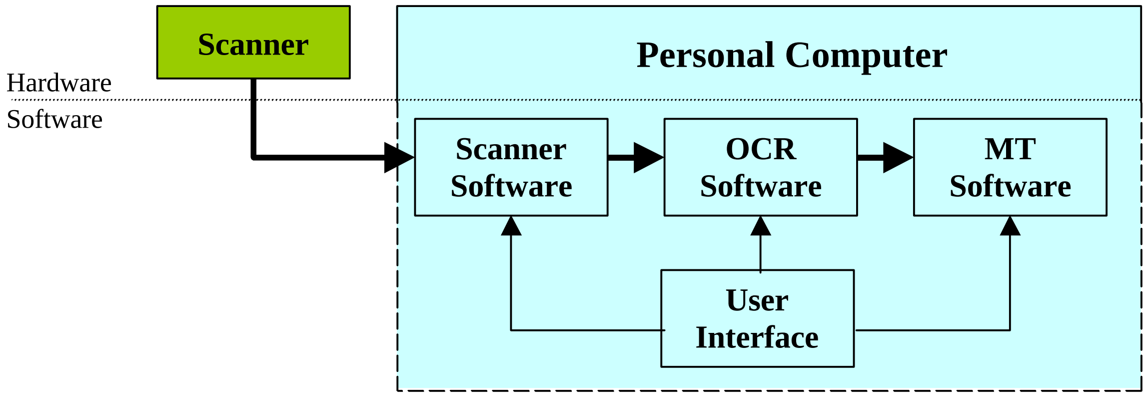


Figure 1: FALCon Process

users were unable to scan these documents using the sheet-fed scanner included as part of the FALCon system. Examples of problem document types included small and/or stiff documents (ID cards), flimsy documents, and documents that were stapled or bound. Soldiers noted that much of the paper being used in Bosnia for printing at that time was of low quality due to the embargo on imports.

Based on this user feedback I proposed the use of a digital camera as an alternative document capture front-end for FALCon. The digital camera provides benefits beyond supporting a wider variety of document types. It can be used to capture text from a variety of targets other than documents. Examples include maps, signs on buildings and walls, and road signs.

Replacing the scanner with a camera can also increase document throughput. Two operators can run the system, one capturing document images with the camera while another processes the images on the computer. The camera can also be used to capture pictures of people and places for reference purposes.

Options to use a flatbed scanner were also considered. One major problem is the size of a flatbed scanner. While new units are currently available in very thin profiles (about 1 inch in some cases), they are still larger then most sheet fed scanners. These units also appear to be more fragile as compared to sheet fed scanners or digital cameras. Flatbed scanners have a lid that must be raised for loading of documents. The hinges used to mount the lid may not hold up to rigorous field use. Also, all units that I have seen to date have some form of manual locking mechanism to protect the scan head during transport. Users would be required to operate this lock prior to equipment transport. Failure to do so would likely result in failure of the flatbed scanner. Earlier tests of hand-held scanners and line scanners also revealed a variety of problems with document image acquisition.

2 Initial Evaluation

The first step in our camera evaluation was to establish initial specifications for the cameras so that we could select units to evaluate. One major specification to establish was the minimum required resolution for the camera. From past experience we have found that low-resolution images such as faxes generally produce poor quality OCR results. When considering the use of a digital camera for document capture we knew that camera resolution would be a major issue. Other specifications considered included size, weight, and cost of the camera, and the interfaces provided for transferring images to the PC.

To support our testing process I established a set of four test documents with ground truth files. Half of these documents are Croatian text in Latinic font and half are Serbian text in Cyrillic font, each in 10 point and 12 point font size. These languages (fonts) were selected for evaluation due to their relevance to the current military presence in the Balkans region. The test documents were assembled from documents used for testing during the integration of the FALCon system.

In order to establish the minimum dots-per-inch (DPI) requirement for the digital camera application we needed to define a process of evaluation independent of the digital camera. To this end we scanned test documents using a flatbed scanner at resolutions from 100 DPI (dots per inch) to 400 DPI. While this method ignores distortions introduced by the digital camera, such as changes in image brightness, optical distortion, and image degradation caused by compression, it does provide a good measure of the "best case" OCR capability for a given DPI value. This in turn establishes a minimum limit on the required camera resolution for a given document size.

In order to test for worst-case pixel rates we scanned the documents at DPI values that were not regularly spaced along the range of resolution values. This was done to reduce the possibility of false results that might occur if we tested only at DPI rates that easily scaled to the values that the software was trained on. The scanned images were then processed to text using the FALCon OCR software and the resulting text files were compared character-by-character to the original source files. The OCR accuracy for the different DPI rates is

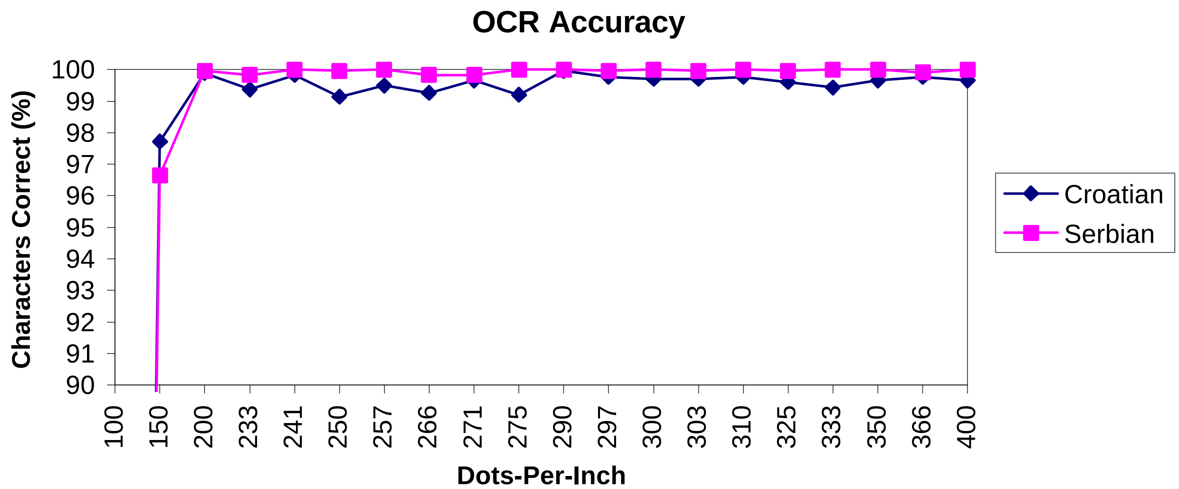


Figure 2: OCR Accuracy versus Document Image Resolution

plotted in Figure 2. This test shows that there is a precipitous decrease in OCR accuracy for DPI rates below 200 DPI.

Having established the need for a minimum of 200 DPI on the document image, we then proceeded to calculate the maximum area that could be spanned using typical high-resolution commercial digital cameras. The resolution for these products is typically 2048 by 1536 pixels. Since the dots-per-inch specification for a scanned document is equivalent to the pixels-per-inch for an image acquired using a digital camera, we calculate the maximum permissible document size as follows:

Maximum Height
2048 pixels * (1 DPI/1 PPI) / 200 DPI = 10.24 inches

Maximum Width
1536 pixels * (1 DPI/1 PPI) / 200 DPI = 7.68 inches.

Since many documents of interest have a 1-inch margin and are roughly 8.5" x 11", these results show that a commercial digital camera with a minimum resolution of 2048 by 1536 pixels should be suitable for this application. Again, remember that this ignores any added image degradation introduced by the camera itself.

With our specifications in hand we performed a market survey collecting information on a wide variety of commercially available digital cameras that had sufficient resolution for the document-processing task. We were able to identify 21 digital cameras that would be suitable for this application. No information was available from manufacturers on possible future cameras with higher resolutions. Assuming that the smallest and lightest camera would be the most portable, and thus the best choice for this application, we selected and purchased the Fuji FinePix 4700 and the Canon Power Shot S20. These cameras were considerably smaller than other products with similar capabilities. While the Fuji FinePix had a higher resolution as compared to many of the other cameras, 4.3M pixels versus 3.3M pixels for the others, we were concerned because this higher resolution is obtained through interpolation. We were unsure what effect this would have, positive or negative, on the final image quality with respect to the operation of the OCR software. Some cameras with resolutions up to 3072 by 2048 pixels were identified in the market survey. These units were not considered for this application due to high cost and the need for operation using a fixed mount such as a stand or tripod.

3 Initial Camera Evaluation

With digital cameras in hand we set out to acquire and evaluate document images. That is when the first problems became apparent. The cameras selected are so small that it becomes difficult to operate them. We found ourselves covering up sensors and flash units with our fingers while trying to take pictures. After a little effort we learned how to hold the cameras in a manner that did not block vital functions. The second problem identified was short battery life. In order to minimize the size of these cameras, manufactures have reduced the size of the battery set. This in turn reduces the time that the camera will operate before the batteries need recharging.

For our camera evaluation we took five pictures of each of the four test documents using both cameras. Initial results from the evaluation of the document images were mixed. On the plus side we found that anticipated problems with the document images being distorted into a keystone shape were minimal. This type of distortion results when the camera is not perpendicular to the image plane of the document. Initial assumptions were that users would find it difficult to adjust the camera to the proper orientation. All of the digital cameras that we evaluated included an

LCD (liquid crystal display) on the digital camera that is used as a through the lens viewfinder. Operators can view the document image on this display during document image acquisition and align the document text to the outside edge of the display. When the camera is in the proper orientation the edges of the document text are aligned with the edges of the display. As a training aid I generated documents with line frames around the outside of the text. During training users can align this frame to the outside edge of the camera display in order to align the camera to the document. After the user understands the process of camera alignment they can then use the edges of the text to align the camera.

On the negative side, evaluation of the document images revealed that the OCR quality was far below what we anticipated. While we had no hard specification on minimum OCR quality we estimated that at least 90% accuracy would be required for the digital camera to be of any use at all. Any errors introduced into the FALCon process affect all remaining steps of the process. If the OCR software recognizes a character in a word incorrectly then the word is misspelled. A misspelled word will either not translate, or the spelling error could result in a real but incorrect word. One way information is lost. The other way false information in provided. From this point of view we needed the best OCR quality that we could achieve. Even with the OCR accuracy at 90% there can still be large numbers of words that do not translate due to spelling errors.

After some initial evaluation we concluded that distortion in the document images was leading to the low OCR accuracy. For document images taken at the extreme limits of the zoom range we found that the lines of the frame surrounding the text area were no longer straight as can be seen in the image shown in figure 3. There was a corresponding distortion of the text near the text frame lines. This problem may be due to either the small size of the lenses on these cameras, or to the automatic settings used with the lenses.

After several rounds of testing in the lab we found that the cameras had “sweet spots” in the zoom adjustment that minimized the optical distortion. Even with this adjustment there were still wide variations in OCR quality across different document images acquired using the same camera settings. For the most part results looked promising.

4 Camera Evaluation

Using the information obtained in the initial camera evaluation we set out to reevaluate our camera selection criteria. A second market survey added one more camera to the list of possible choices. We relaxed the requirement for minimum camera size and started looking for information on the battery type and quantity and the operating time on a set of batteries. Optical image distortion due to zoom adjustment was a much

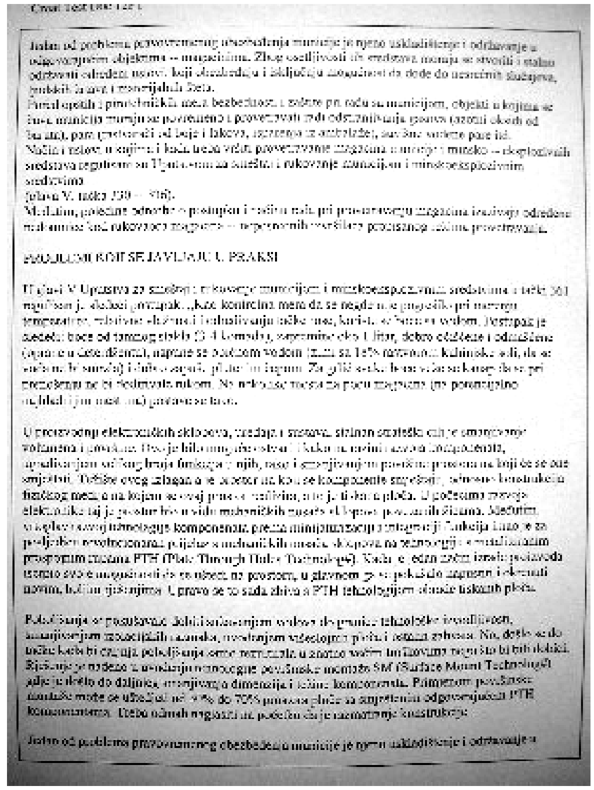


Figure 3: Document Image at Wide Zoom Setting

harder parameter to characterize. After some thought I spent an afternoon at a local electronics store where a salesman was kind enough to show me every high-resolution digital camera that they carried. I used a sheet of paper as a target and pre-evaluated all of the cameras in an attempt to identify those cameras with the least optical distortion. Back at the lab I correlated the list of best optical choices against the size, cost, and battery parameters of the cameras. I then selected and purchased a Ricoh RDC-7, Casio QV-3000EX, and Sony Cyber-Shot DSC-S70, and borrowed a Nikon CoolPix 990.

With the new cameras in the lab we processed a limited number of documents on each camera for different zoom settings in order to get a baseline on optical quality. Again we found that most of the cameras had particular zoom settings that produced good OCR results. In order to minimize the camera test set I decided that we would process document images for each camera at zoom settings of wide, middle, and narrow. We captured and evaluated five images of each of the four test documents for all of the cameras. With six cameras, four documents, three zoom settings, and five pictures per document we captured a total of 360 images.

This second round of camera testing revealed several new features of the digital cameras. For four of the six cameras the lens assembly protrudes from the camera body during operation, and retracts into the camera body when the camera is turned off. This provides a

Table 1: OCR Accuracy for Digital Cameras

		Document Type/Font Size (% Accuracy averaged over 5 documents)			
Camera	**Zoom Setting**	**Croatian 10 pt**	**Croatian 12 pt**	**Serbian 10 pt**	**Serbian 12 pt**
Canon	Wide	65.8%	79.9%	68.7%	70.7%
	Middle	72.4%	77.6%	68.9%	81.8%
	Narrow	70.3%	81.7%	54.3%	92.5%
Casio	Wide	78.3%	53.8%	47.5%	76.0%
	Middle	94.8%	88.6%	81.9%	74.5%
	Narrow	38.1%	56.1%	41.5%	85.4%
Fuji	Wide	54.5%	67.5%	46.6%	72.9%
	Middle	82.6%	90.9%	86.0%	98.4%
	Narrow	58.1%	51.3%	53.9%	70.0%
Nikon	Wide	73.1%	72.9%	76.6%	76.4%
	Middle	99.3%	99.4%	97.9%	98.2%
	Narrow	72.0%	75.0%	38.7%	71.3%
Ricoh	Wide	92.1%	79.2%	85.0%	64.7%
	Middle	93.2%	98.9%	92.2%	96.1%
	Narrow	78.7%	53.9%	83.6%	68.0%
Sony	Wide	80.8%	78.1%	84.2%	91.0%
	Middle	98.4%	99.0%	96.8%	99.7%
	Narrow	96.9%	95.4%	96.7%	99.9%

potential path for dirt to enter the camera. Further, if the lens protrudes from the camera body then it may be damaged in field use.

Several of the cameras use custom battery packs that would not be readily available for field replacement. In my limited experience with soldiers in field exercises I found that AA batteries were always available. Those cameras that use AA batteries would then be more desirable for this application because soldiers could find replacement batteries when needed.

While all of the cameras included a USB (universal serial bus) connection for image transfer, I found it easier to simply remove the image storage card and mount it in the PC using a PC-Card adapter. This worked well for Compact Flash memory and for the Sony Memory Stick. The Smart Media card used by the Fuji and Ricoh cameras proved difficult to handle in this operation.

After capturing the document images we converted them to text using the OCR software and compared the resulting text files to the ground truth documents to measure the OCR accuracy. All changes in page format were ignored because they do not change the content of the document. The results of this testing, averaged for each set of five test documents, are shown in Table 1. Most of the cameras worked best with the zoom set to the middle of the range of adjustment. Several of the cameras had greater than 90% OCR accuracy for this case, and the Sony and Nikon cameras had greater than 96% accuracy. Also, most of the cameras suffered severe degradation in OCR quality when the zoom was set to the wide or narrow setting. The Sony camera was the exception in the narrow zoom setting. For most zoom settings the OCR quality for the 12-point text is better than that for the 10-point text. This result was anticipated because the 12-point text has more pixels on each character. This in turn provides more information to the OCR software for recognition. Exceptions occurred mostly in cases where the OCR quality was below 90%. Again the Sony camera was the exception.

As in the initial testing we found that OCR results for some of the cameras varied considerably from image to image for the same document with the same zoom, flash, and white balance settings. This variation is illustrated by the data shown in Table 2, which lists the OCR error counts for each of the five document images taken at each zoom setting for the Canon camera. To better understand this problem we repeated the testing process on some of the cameras and obtained similar results. We then examined the document images using image-processing software. Using only visual inspection we found minimal variation between images that had good OCR results and images that had bad OCR results. After some processing we found that many of the images with bad OCR results had substantial variations in illumination across page. Figure 4 shows a document image that has been posterized, or converted to a fixed number of image intensity levels, in order to show the variation in illumination.

Assuming that we could attribute some of the OCR error to improper document illumination we proceeded to adjust the flash level and image brightness on the cameras with those options in order to try to optimize image quality. We repeated the image acquisition and evaluation process for selected zoom settings on selected cameras. In some cases these adjustments improved OCR accuracy while in others they decreased it. Also, the number of OCR errors remained more consistent from image to image, even for those cases

Table 2: OCR Character Error Counts for Canon Camera

	OCR Character Error Count				
Zoom	**Image**	**Croatian 10 pt**	**Croatian 12 pt**	**Serbian 10 pt**	**Serbian 12 pt**
Wide	1	2286	595	1046	1588
	2	1602	855	2149	501
	3	1791	608	1024	1313
	4	1385	732	1104	894
	5	1736	749	2303	872
Middle	1	3475	689	1037	220
	2	1129	2548	791	608
	3	479	404	2330	1381
	4	800	192	921	696
	5	1218	109	2516	298
Narrow	1	1924	83	2501	611
	2	2696	260	2652	611
	3	840	1082	893	120
	4	769	414	2475	59
	5	1432	1389	2634	120

where OCR accuracy was not improved. Unfortunately, it would not be reasonable to expect users in the field to make a variety of adjustments to the digital camera in order to obtain suitable results. We needed other options.

One option considered was to correct the image intensity by thresholding. With thresholding the pixels of the image are converted to either the maximum or minimum possible value based on their value as compared to the threshold level. Figure 5(a) shows a document image that has been processed in this manner. Note that the corners of the image are darkened. This occurs because the threshold level was set to obtain good text clarity at the center of the page. If the threshold is set low enough to remove the corner darkening then the text at the center of the page will be washed out. Both results decrease the OCR accuracy.

A coworker suggested correcting for the flash variation by performing a scaling correction proportional to the distance from the center of the image. In order to test this possibility I used two commercial software packages to generate a correction image that closely approximated the inverse of the lighting pattern on the document images. I then normalized one of the document images using the correction image. Evaluation of the resulting image showed a substantial increase in OCR accuracy. For our case we found that the center of the camera flash pattern moved on the page as the distance from the camera to the document changed. Changes in the distance from the camera to the document also resulted in different rates of change of the illumination across the page. This would require the illumination correction across the page to be calculated for each document image processed. While it sounds like a good idea this process does not seem feasible for our application.

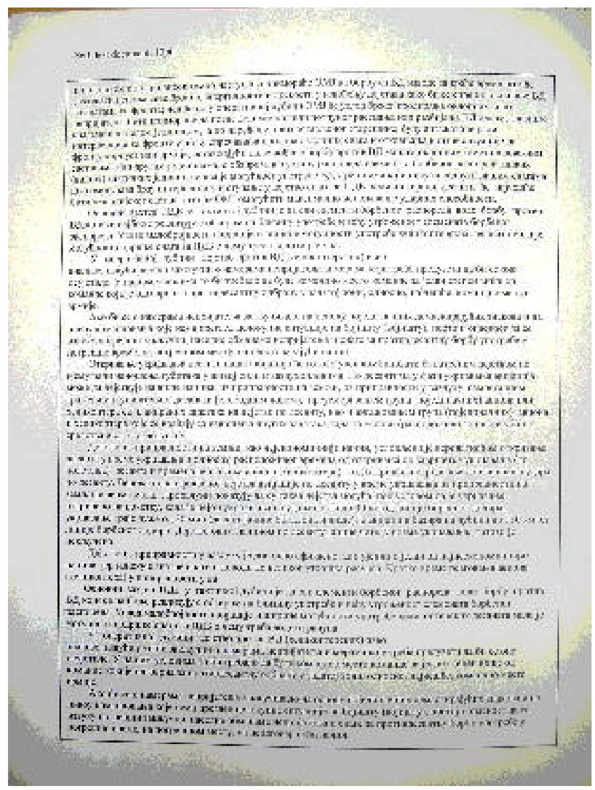
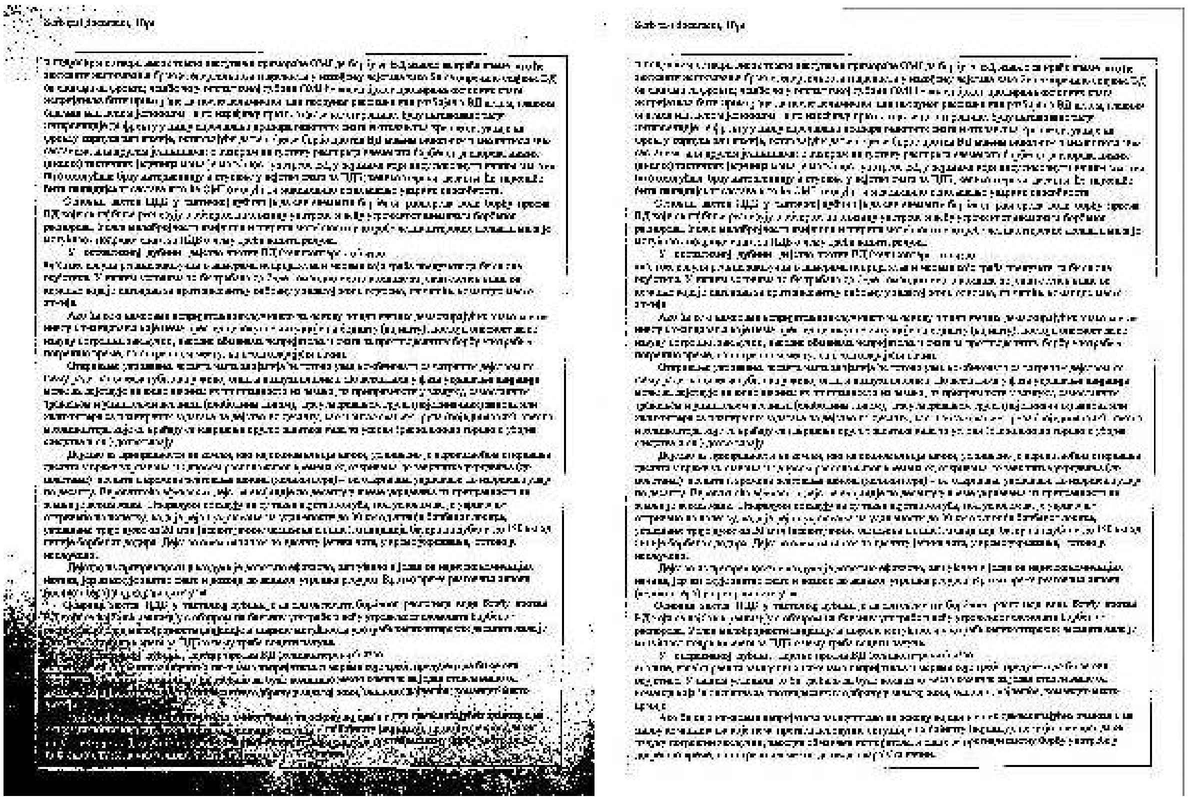


Figure 4: Posterized Document Image Showing Changes in Illumination

As an alternative I proposed performing a threshold process on sub-blocks of the image. By dividing the image into sub-blocks we minimize the change in image illumination across the area to be processed and eliminate the problem of text in one part of the image being lighter than paper in another part of the image. In order to prove this concept I manually processed an image by thresholding sub-blocks of the image. This process substantially reduced the OCR error.



(a) (b)

Figure 5: (a) Full-Page Threshold Document Image, (b) Sub-Block Threshold Image

Figure 5 shows the full-page threshold image on the left and the sub-block threshold image on the right. An algorithm was then developed to perform the sub-block threshold automatically. One interesting aspect of the sub-block threshold process was the selection of the number of sub-blocks into which the image would be divided. I selected the number of sub-blocks to process by working with a worst case image obtained by taking a picture of a document placed against a page in an open book. This document image was severely distorted both optically and in image intensity due to the curving of the book pages by the book binding. After several processing steps I settled on 256 sub-blocks in a 16 by 16 block pattern.

Using the sub-block threshold algorithm we reprocessed the document images acquired using the Nikon and Casio cameras. The OCR accuracy results are shown in table 3. Again these are averages of the results over 5 document images. Note that for most cases the OCR accuracy has increased to greater than 90%. We also found that the resulting OCR character errors were relatively consistent from image to image for a given document, camera, and zoom setting. Any remaining error is most likely attributable to the optical distortion caused by the camera lens. Using the sub-block threshold we found similar results with a partial test of the document images from the other cameras.

As part of this process we identified another interesting feature of the digital camera images. Most of these cameras store their images in JPEG format with no reference for the image dimensions. This makes sense because the camera does not know how large the target object is. The OCR software used with FALCon interprets these camera images as very large documents at 72 pixels-per-inch. This causes two problems. First, the text output from the OCR has a very large font. Second, the standard FALCon process cannot pass the image. We have modified the sub-block threshold

Table 3: OCR Accuracy for Sub-Block Threshold Document Images

		% Accuracy/% Improvement (averaged over 5 documents)			
Camera	**Zoom**	**Croatian 10**	**Croatian 12**	**Serbian 10**	**Serbian 12**
Nikon	Wide	84.7%/+11.6%	90.6%/+17.7%	90.3%/+13.7%	98.6%/+22.2%
	Middle	99.3%/0%	99.5%/+0.1%	99.6%/+1.7%	99.9%/+1.7%
	Narrow	97.5%/+25.5%	98.8%/+23.8%	98.9%/+60.2%	99.8%/+28.5%
Casio	Wide	91.9%/+13.6%	84.3%/+30.5%	95.5%/+48%	92.7%/+16.7%
	Middle	99.3%/+4.5%	99.4%/+10.8%	99.3%/+17.4%	99.9%/+25.4%
	Narrow	90.2%/+52.1%	93.8%/+37.7%	97.1%/+55.6%	98.0%/+12.6%

process so that it sets the pixels-per-inch setting to 200 in order to correct this problem.

5 Alternative Lenses

Since the camera optics appeared to affect the OCR process we considered possible options to improve the quality of the optics. I ordered a close up lens for the Nikon camera for evaluation. I selected the Nikon for this test because it is the only unit being evaluated that has a threaded mount to accept accessory lenses. After receiving the lens we mounted it on the camera and started looking at documents. Unfortunately, the lens blocks the flash output making it unusable in our application. Also, while the optical distortion was reduced in some areas of the image, it was not reduced in others. Overall I concluded that the close up lens was not suitable for this application.

6 Camera Selection

With the development of the sub-block threshold process we demonstrated the capability to compensate for OCR errors resulting from variation in document image brightness. Table 3 shows that two of the cameras have greater than 90% OCR accuracy for the narrow zoom setting and greater than 99.3% accuracy for the middle zoom setting. A partial evaluation of document images from the other four cameras evaluated yielded similar results. We were unable to reevaluate all of the document images for all cameras prior to the final camera selection or the writing of this paper due to the loss of support staff.

With the OCR accuracy testing complete we needed only to make our final selection of a digital camera for our application. The initial selection criterion was to get the best possible OCR accuracy. After working with the digital cameras in the lab we placed a much greater emphasis on the need to have a camera that was easy to operate and that we considered would have minimal problems in the field. Table 4 lists the camera specifications for the units that we evaluated.

Based on our experience in the lab we reviewed the camera size specifications and decided that selecting the smallest cameras would not be the best option. If the smallest cameras were difficult to use it the lab the problem could only be worse in a military field environment.

Next we eliminated some of the camera by looking at the requirement for field use. We eliminated those cameras with lenses that protrude from the camera body during operation. This would reduce the possibility of camera failures in the field due to dirt infiltration and physical impact to the lens assembly. With two cameras remaining, the Ricoh and Nikon, we looked at the battery and memory card types. The Nikon uses the Compact Flash card and AA-batteries, both pluses for field use. The Ricoh uses the Smart Media card and a Lithium-ion battery pack. We found the Smart Media card difficult to handle in our camera evaluation, and as mentioned previously it would probably be difficult to locate a replacement Lithium-ion battery pack in the field.

Camera cost was a consideration in the selection of our cameras for evaluation. Several very high-resolution cameras were not considered due to the high cost. For the six cameras evaluated the variation in cost was small. As a result cost was not considered in the final camera selection.

Based on our updated selection criteria and the results of the OCR evaluation we selected the Nikon CoolPix 990 as our camera of choice for this application.

7 User Evaluations

Soldiers of the U.S. Army will test the digital camera extension for FALCon in Advanced Concepts Technology Demonstration (ACTD) field exercises in 2001. ARL staff will train users in the application of the digital camera extension for FALCon. We hope to get user feedback on the viability of this concept along with suggested system improvements

As a preliminary user test I had one of my coworkers, picked in part because he was not a camera expert, use the Nikon digital camera to acquire document images. I provided five minutes of training, explaining how to operate the camera and how to take pictures of the documents. With that my coworker was on his own to acquire five document images for each of the four test documents. The OCR accuracy for these document images was only 1% lower than results that I obtained after hours of practice capturing document images in the lab.

8 Conclusion

While user evaluation has not yet been performed, it is

Table 4: Camera Specifications

Camera	Size (cu. in.)	Weight	Lens Pop Out	Memory Type	Cost	Battery Type
Casio	36.15	0.7 lbs.	Yes	Compact Flash	$999	AA x 4, Ni-MH
Canon	14.39	0.7 lbs.	Yes	Compact Flash	$1030	Ni-MH pack
Fuji	15.31	0.56 lbs.	Yes	Smart Memory	$999	AA x 2, Ni-MH
Nikon	27.44	0.81 lbs.	No	Compact Flash	$999	AA x 4, Ni-MH
Ricoh	15.37	0.59 lbs.	No	Smart Memory	*$800	Li-ion pack
Sony	31.82	0.96 lbs.	Yes	Memory Stick	$899	Li-ion pack

* = Price quoted after other products, prices decrease over time

apparent that high-end commercial digital cameras can be used in a lab environment to capture document images for processing purposes. We can capture an 8.5” x 11” text document with one image using a 3.3-mega pixel digital camera and obtain reasonable OCR results. Future increases in digital camera resolution will lead to better OCR results for this application.

Given the proper set of adjustments, most of the digital cameras we evaluated appear capable of supporting document capture for evaluation using the FALCon system. The zoom settings for these cameras can be adjusted to minimize optical distortion. Unfortunately, some of these adjustments make it difficult to acquire document images by placing the target document far from the user. The distance to the document amplifies any slight motion on the part of the user while taking the document image, possibly resulting in image blur and improper framing of the document.

Many of the document images captured using the digital cameras that we evaluated suffered some degradation of the OCR accuracy due to changes in illumination across the image. We were able to compensate for this by running a sub-block threshold process on the document images. This sub-block threshold process resulted in substantial improvement in OCR accuracy.

For the digital camera extension for FALCon we selected the Nikon CoolPix 990 based on our evaluation of the Nikon’s perceived usability in a military field environment, and on our evaluation of OCR accuracy for digital cameras.

9 Acknowledgements

We would like to express our sincere gratitude to the Human Intelligence Support Tools Advanced Concepts Technology Demonstration (ACTD) staff for funding this effort. I would also like to thank Richard Chang, a Computer Science student from the University of Maryland, for assisting in the data collection and evaluation.